

REMARKS

U.S. Patent No. 6,286,140, to Ivanyi ("the '140 patent"), is directed to a signal receiving device which, upon the occurrence of a pre-specified event, reads the data from each of the monitored devices embedded therein, including a real-time event clock, thereby recording the post-event operational state of the signal receiving device. The above-described data is stored in a database residing in the signal receiving device until such time as a central processing computer polls the signal receiving device to initiate a data transmission or data upload. Once received by the central computer, the information is stored in a central database, along with information regarding the demographics of the viewer, subscriber, or customer, as well as other data or information that may be compiled with viewer, subscriber, or customer consent or permission. The '140 patent suggests that data stored in the central database can be analyzed to determine viewer behavior or viewer responses to various programming and advertising subject matter, but does not teach a means for such analysis. The '140 patent further suggests that such analysis can facilitate assessment of the effectiveness of programs and advertising commercials that, thereafter, may be made available to advertisers, but does not teach or suggest a means for such analysis.

Applicant's invention is patentable over the '140 patent. Applicant's invention is generally directed to a market data acquisition system and method that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. In one aspect of the invention, certain set-top box events may be invalidated based on past set-top box operator behavior. Set-top box data is correlated with content attributes and demographic information for a region in which the set-top box is located. Such correlations are drawn using techniques which reduce the effect of sampling error and sample bias and which increase correlation result dataset specificity. The correlations result in data indicating which content was experienced by one or more demographic groups, which can be used to determine the relative rating of such content. Another aspect of Applicant's invention is the determination of individual set-top box user characteristics. This can be especially useful when set-top box data is collected in a privacy-compliant manner, as is preferred in the Applicant's invention. A further aspect of the Applicant's invention is the determination of the effect of content attributes on content ratings. The '140 patent does not teach or enable Applicant's

claimed invention. The Court of Appeals for the Federal Circuit has consistently held that "Anticipation requires the presence in a single prior art reference disclosure of each and every element of the claimed invention, arranged as in the claim."

Lindemann Maschinenfabrik GmbH v. American Hoist & Derrick, 221 USPQ 481, 485 (Fed. Cir. 1984). The '140 patent clearly fails to disclose structure positively recited and claimed in applicant's independent claims.

U.S. Patent No. 6,286,005, to Mark E. Cannon ("the '005 patent"), teaches a method of storing and accessing television viewership information provided by a third party, in which incoming data is converted into an unusual database file format. The '005 patent is directed to receiving, on a weekly basis, large quantities of data and converting such data into an appropriate format. Once in this form, the converted data can be moved from location to location; organized and analyzed in computer memory; filtered based on demographic criteria; and can facilitate retrieval of information for a sample group across multiple weeks.

Applicant's invention is patentable over the '005 patent. Applicant's invention is generally directed to a market data acquisition system and method that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. The '005 patent clearly fails to disclose structure positively recited and claimed in applicant's independent claims.

U.S. Patent No. 6,029,176, to Mark E. Cannon ("the '176 patent"), teaches a method and apparatus for quickly reviewing, manipulating, and analyzing large quantities of computer-based data relevant to television-viewing consumers. The '176 patent is directed to the analysis of data supplied by the A. C. Nielsen Company based on television viewing logs from specialized equipment attached to televisions in homes. The '176 patent teaches a method of converting the A. C. Nielsen Company data to a proprietary format, and filtering techniques that take advantage of the proprietary data format. The '176 patent also teaches the use of a graphical user interface to provide access to the data stored in the proprietary format.

Applicant's invention is patentable over the '176 patent. Applicant's invention is generally directed to a market data acquisition system that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. The '176 patent does not teach or suggest the use of set-top event data. It is well-established that, in order to show obviousness,

all limitations in the claim must be taught or suggested by the prior art. In Re Boyka, 490 F.2d 981, 180 USPQ 580 (CCPA 1974); MPEP § 2143.03. It is error to ignore specific limitations distinguishing over the references. In Re Boe, 184 U.S.P.Q. 38, 505 F.2d 1297 (C.C.P.A. 1974); In Re Saether, 181 U.S.P.Q. 36, 492 F.2d 849 (C.C.P.A. 1974); In Re Glass, 176 U.S.P.Q. 489, 472 F.2d 1388 (C.C.P.A. 1973). The '176 patent clearly does not disclose all the limitations in Applicant's claims.

U.S. Patent No. 6,289,514, to Link ("the '514 patent"), describes a system and method for identifying television programming, identifying and capturing consumer television program viewership behavior and providing real-time reporting of that information to interested parties, while also providing verification of actual delivery of advertising and/or program content. The '514 patent teaches that it is important to establish a database of "cluster codes", or mathematical groupings of the viewing population such that there is a demographic segmentation of the viewing population based on socioeconomic factors. The cluster code database should also include a merger of possible clustering codes with customers' actual addresses from a cable customer database and cross-referencing that to a set top box ID data base. STB events are aggregated by time, channel, cluster code and head end. The '514 patent teaches that reports consisting of household ratings can be generated from the aggregated event data in near-real time.

Applicant's invention is patentable over the '514 patent. Applicant's invention is generally directed to a market data acquisition system that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. Applicant's invention does not rely on cluster codes when creating such statistics. Instead, the statistics provided by Applicant's invention are based on demographic information. The Court of Appeals for the Federal Circuit has consistently held that "Anticipation requires the presence in a single prior art reference disclosure of each and every element of the claimed invention, arranged as in the claim." Lindemann Maschinenfabrik GmbH v. American Hoist & Derrick, 221 USPQ 481, 485 (Fed. Cir. 1984). The '514 patent clearly fails to disclose structure positively recited and claimed in applicant's independent claims.

U.S. Patent No. 6,020,883, to Herz, et al. ("the '883 patent") teaches a system and method for calculating which program characteristics are preferred by a customer to give a measure of how well that customer should like a given program. The '883 patent does not teach or suggest Applicant's invention. Applicant's invention is

generally directed to a market data acquisition system that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. Applicant's invention does not predict how well a specific individual will like a given program.

U.S. Patent No. 6,088,722, to Herz, et al. ("the '722 patent") teaches a system and method for calculating which program characteristics are preferred by a customer to give a measure of how well that customer should like a given program, and provides a feedback mechanism to determine whether the customer actually liked the program. The '722 patent does not teach or suggest Applicant's invention. Applicant's invention is generally directed to a market data acquisition system that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. Applicant's invention does not predict how well a specific individual will like a given program.

U.S. Patent No. 6,298,348, to Eldering ("the '348 patent") teaches a system which learns the buying habits of consumers based on data received from a point of purchase, such as a register in a grocery store. The '348 patent teaches that such point of purchase data can be used to develop consumer demographic profiles using a vector-based representation of the probability that a consumer falls within a certain demographic category such as an age group, gender, household size, or income range. The '348 patent also teaches that a consumer's product preferences can also be discerned from the learned buying habits to determine which products a consumer is likely to purchase in the future. The demographic profile and product preferences are updated with each purchase. The updating process uses a weighting factor which determines the importance of the purchased product with respect to all of the products purchased in a particular product category. The '348 patent also teaches that print advertisements can then be targeted to a consumer based on the consumer's demographic profile and product preferences.

Applicant's invention is patentable over the '348 patent. Applicant's invention is generally directed to a market data acquisition system that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. The '348 patent does not disclose all the limitations in Applicant's claims, thus Applicant's invention is patentable over the '348 patent.

U.S. Patent No. 6,216,129, to Eldering (“the ‘129 patent”) teaches an advertisement selection system in which vectors describing an actual or hypothetical market for a product or desired viewing audience can be determined. A consumer characterization vector is correlated with an ad characterization vector to determine the suitability of the advertisement to a specific consumer. The consumer characterization vector describes statistical information regarding the demographics and product purchase preferences of a consumer, and is developed from previous purchases or viewing habits.

Applicant’s invention is patentable over the ‘129 patent. Applicant’s invention is generally directed to a market data acquisition system that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. The ‘129 patent does not disclose all the limitations in Applicant’s claims, thus Applicant’s invention is patentable over the ‘129 patent.

U.S. Patent No. 5,872,588, to Aras, et al. (“the ‘588 patent”), teaches a method and apparatus through which audio-visual materials are encoded with special content codes. The content codes allow a home station to collect information on the subscribers selection of audio visual material streamed to the home station and to collect the encoded content codes. The collected codes are then sent to collection centers for processing.

Applicant’s invention is patentable over the ‘588 patent. Applicant’s invention is generally directed to a market data acquisition system that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. The ‘588 patent does not disclose all the limitations in Applicant’s claims, thus Applicant’s invention is patentable over the ‘588 patent.

U.S. Patent Application Publication No.: 2001/0049620, to Blasko (“the ‘620 publication”), teaches a system and method for transaction profiling in a privacy-protected manner based on transaction data, wherein the transaction generally refers to an intentional action by a user. The ‘620 publication teaches that transaction data relates to programming and advertisements watched by the user over a predetermined period of time. Transaction profile vectors are computed based on the transaction data, wherein the transaction profile vector may include demographic attributes such

as age, household size, income level of the user, or preference attributes indicating probable products and services preferred by the user. The '620 publication teaches that the transaction profile vector preferably takes place local to the transaction.

Applicant's invention is patentable over the '620 publication. Applicant's invention is generally directed to a market data acquisition system that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. The '620 publication does not disclose all the limitations in Applicant's claims, thus Applicant's invention is patentable over the '620 publication. In addition, applicant believes that the '620 publication is not valid prior art under 35 U.S.C. §102(a) because Applicant's filing and priority dates predate those of the '620 publication. Applicant also believes that the '620 publication is not valid prior art under 35 U.S.C. §102(b) or 35 U.S.C. §102(e) because the publication date for the '620 publication is well after Applicant's filing date. Further, Applicant believes that it would be improper to suggest that the '620 publication could be combined with other references and thus to reject Applicant's claims under 35 U.S.C. §103(a) because Applicant's patent application has a priority date predating that of the '620 application.

U.S. Patent Publication No.: 2001/0021994, to Nash ("the '994 publication"), teaches a television system which enables advertisements to be targeted at viewers who have a particular interest in the products or services being promoted by the advertisement which comprises reviews from a plurality of reviewers commissioned by an advertiser and alternatively or additionally by independent reviewers. The reviews are encoded into a data channel associated with the advertisement. A product rating decoder is provided in a receiver that decodes the data in the data channel and selects advertisements for display based on the data and a user profile generated either explicitly by the user entering preferences via a user interface or implicitly by monitoring the type of program selected for viewing by the user.

Applicant's invention is patentable over the '994 publication. Applicant's invention is generally directed to a market data acquisition system that correlates set-top box event data with content attributes and user demographic information to calculate content experiencing statistics. The '944 publication does not disclose all the limitations in Applicant's claims, thus Applicant's invention is patentable over the '994 publication. In addition, applicant believes that the '994 publication is not valid prior art under 35 U.S.C. §102(a) because Applicant's filing and priority dates predate

those of the '994 publication. Applicant also believes that the '994 publication is not valid prior art under 35 U.S.C. §102(b) or 35 U.S.C. §102(e) because the publication date for the '620 publication is well after Applicant's filing date. Further, Applicant believes that it would be improper to suggest that the '994 publication could be combined with other references and thus to reject Applicant's claims under 35 U.S.C. §103(a) because Applicant's patent application has a priority date predating that of the '994 application.

**PRELIMINARY AMENDMENT UNDER 37 C.F.R. §1.111**

Prior to substantive examination in the United States Patent and Trademark Office, please cancel Claims 32-34, 35, 40 through 62, 65, and 66 without prejudice toward the filing of one or more additional patent applications directed toward the content thereof. In addition, please execute the following amendments to the Claims:

- In Claim 1, strike the words "or set of demographic groups" from the end of the final element of the Claim and insert "t least one" between "experienced by a" and "demographic group" in the same element. For the Examiner's convenience, attached hereto is a copy of the Claims as amended and a copy illustrating changes made to the Claims.
- In Claim 6, strike the words "present invention" at the end of the Claim and insert the word "system" therefor.
- In Claim 7, insert the phrase "as part of the system" at the end of the Claim.
- In Claim 10, strike the word "the" from between the words "result" and "generation" and insert the word "of" between the words "generation" and "reports".
- In Claim 11, insert the word "for" between the words "generated" and "individual".
- In Claim 67, strike the phrase "such as" from the preamble and insert the word "comprising" therefor; strike the word "such" from the first element and insert the phrase "data relationship" therefor.
- In Claim 71, strike the "20" and insert "67" therefor.

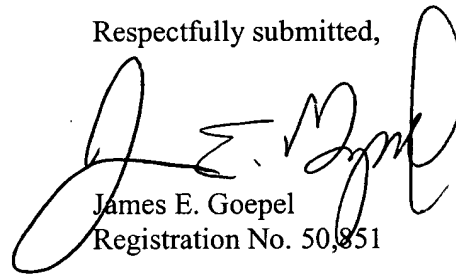
REMARKS

Claims 1-33, 36-39, 63, 64, and 67-80 are presently under consideration. Attached hereto on the pages captioned "Version with Markings Illustrating Changes" are marked-up copies of the claims which illustrate the changes made thereto. In addition, attached hereto on the pages captioned "Claims as Amended" are copies of the claims as amended.

CONCLUSION

Should the Examiner feel that there are any issues outstanding after consideration of this response, the Examiner is invited to contact Applicants' undersigned representative to expedite the prosecution. You are hereby authorized to charge or credit any deficiency or overpayment to our Deposit Account No. 50-0653.

Respectfully submitted,



James E. Goepel  
Registration No. 50,851

Greenberg Traurig  
1750 Tysons Boulevard, 12<sup>th</sup> Floor  
McLean, Virginia 22102  
703-903-7533  
Filed: August 15, 2001

\\TCO-srv01\GOEPEL\102797v01\27BH011.DOC3/18/02\37865.010200



We claim as our invention:

1. (Amended) A market data acquisition system, comprising:
  - a means for retrieving event and embedded content data from a plurality of set-top boxes;
  - a means for retrieving content attributes from a content attribute database;
  - a means for correlating retrieved set-top box event data with content attributes to produce data indicating which content was experienced through the plurality of set-top boxes;
  - a means for retrieving demographic information from a demographic information database; and
  - a means for correlating demographic information to data indicating which content was experienced through the plurality of set-top boxes to produce, in response to a query, data indicating content experienced by at least one demographic group [or set of demographic groups].
2. The market data acquisition system of Claim 1, in which said state-change data collection means collects data from said set-top boxes without access to set-top box specific personal or demographic information, thereby providing a layer of privacy to set-top box assignees.
3. The market data acquisition system of Claim 2, in which set-top box specific demographic or other personal data may be collected when requested or with approval given by a set-top box assignee, governmental agency, or other such authority.
4. The market data acquisition system of Claim 3, in which a list of set-top box identification numbers and zip codes or other geographic identifiers corresponding to set-top box installation points is provided to the present invention for each set-top box.
5. The market data acquisition system of Claim 1, in which said content attribute database is maintained as part of the system.
6. (Amended) The market data acquisition system of Claim 1, in which said content presentation system is maintained external to the system [present invention].
7. (Amended) The market data acquisition system of Claim 1, in which said demographic information database is maintained as part of the system.
8. The market data acquisition system of Claim 1, in which said demographic information database is maintained externally.

9. The market data acquisition system of Claim 1, in which said queries are entered through a graphical, command-line, or natural language interface.
10. (Amended) The market data acquisition system of Claim 9, in which said queries can result in [the] generation of reports for any time segment or set of time segments with high precision.
11. (Amended) The market data acquisition system of Claim 9, in which said queries result in the generation of reports generated for individual content or for a set of content.
12. The market data acquisition system of Claim 9, in which said queries result in generation of said reports for persons fitting a demographic specification, persons fitting a demographic category, or persons fitting sets of demographic specifications and demographic categories.
13. The market data acquisition system of Claim 9, in which said queries result in reports generated for specific behaviors.
14. The market data acquisition system of Claim 9, in which said queries include one or more highly-specific times, demographic specifications, viewer behaviors, and content descriptions.
15. The market data acquisition system of Claim 9, in which said results are presented in a graphical manner, such as through a pie chart or bar graph.
16. The market data acquisition system of Claim 9, in which said results are presented as a spreadsheet or other grid.
17. The market data acquisition system of Claim 9, in which said results are presented as natural language.
18. The market data acquisition system of Claim 1, in which said content information is obtained from a source external to the present invention.
19. The market data acquisition system of Claim 1, in which said content information is embedded in content as it is presented to a set-top box.
20. A method of correlating dynamic and static datasets sharing at least one common characteristic and having an assumed relationship, and using such correlations to determine rule systems between the sets, comprising the steps of:

- selecting subsets of said datasets sharing a common characteristic;
  - expressing the assumed relationship as a mathematical assumption;
  - defining an error function which describes the two datasets in terms of said mathematical assumption;
  - performing fitting procedures to account for errors in the assumed relationship;
  - and
  - performing fitting procedures which account for errors in the definition of the common subsets.
21. The method of Claim 20, in which said dynamic data corresponds to set-top box event data.
22. The method of Claim 21, in which said static data corresponds to demographic data.
23. The method of Claim 22, in which correlations are drawn between set-top box event data and demographic to determine the relationship of demographics to content viewership.
24. A method of testing assumptions pertaining to relationships between two disparate datasets sharing at least one common aspect, comprising the steps of:
- entering such assumptions through a user interface;
  - selecting sample data from a first dataset;
  - determining correlations between said selected data and data stored in a second dataset; and
  - establishing assumption validity based on such correlations.
25. A method of determining individual characteristics by correlating dynamic and static datasets sharing at least one common characteristic and having an assumed relationship, comprising the steps of:
- selecting subsets of said datasets sharing a common characteristic;
  - expressing the assumed relationship as a mathematical assumption;
  - defining an error function which describes the two datasets in terms of said mathematical assumption;
  - performing fitting procedures to account for errors in the assumed relationship;
  - storing such correlations in an individual-specific array; and
  - iteratively repeating this process.
26. The method of Claim 25, in which said dynamic dataset corresponds to set-top box data.

27. The method of Claim 26, in which said static dataset corresponds to demographic data.

28. The method of Claim 27, in which said individual-specific data corresponds to a set-top box identification number or other privacy-compliant identification number.

29. The method of Claim 28, in which an IDM algorithm determines said correlations.

30. A method of dynamically determining the demographic identity of an individual operating a set-top box, comprising the steps of:

monitoring set-top box events for a plurality of set-top boxes;

correlating set-top box events with demographic characteristics;

applying IDM calculation techniques to determine probabilities for demographic characteristic and set-top box event dataset correlations;

ascribing demographic characteristic probabilities to each set-top box over time based on observed set-top box events and their relationship to such IDM probabilities;

evaluating such ascribed demographic characteristic probabilities over time through statistical analysis;

fitting probabilities ascribed to demographic characteristics to statistically determine the most likely set of constant dataset possibilities for each set-top box; and,

fitting set-top box possibility sets to IDM probability sets for a set-top box event.

31. The method for determining the demographic identities of individuals in a home, business, or other location containing a set-top box according to the method of Claim 30, further comprising the steps of:

storing said demographic identities in an array over time; and

applying statistical analyses to said array to determine predominant demographic identities for a given set-top box.

32. (Cancelled)

33. (Cancelled)

34. (Cancelled)

35. (Cancelled)

36. A method of determining the effect of content attributes on content ratings, comprising the steps of:

obtaining content attributes from embedded content information or from external sources;

**Version with Markings Illustrating Changes**

recording set-top box events as content is experienced;  
correlating set-top box events to content attributes; and,  
analyzing such correlations over time to determine the effect of content attributes on content ratings.

37. The method of Claim 36 in which said content attributes include times at which various content attributes are presented to a set-top box, thereby allowing the present invention to provide detailed correlations between such attributes and set-top box events.

38. A method of determining the effect of content attributes on content ratings for a specific demographic group, comprising the steps of:

obtaining content attributes from embedded content information or from external sources;  
recording set-top box events as content is experienced;  
correlating set-top box events to content attributes;  
correlating set-top box events and content attributes to demographic characteristics for each set-top box; and  
analyzing such correlations over time to determine the effect of content attributes on content ratings for specific demographic groups.

39. The method of Claim 38 in which said content attributes include times at which various content attributes are presented to a set-top box, thereby allowing the present invention to provide detailed correlations between set-top box events, set-top box demographics, and content attributes.

40. (Cancelled)

41. (Cancelled)

42. (Cancelled)

43. (Cancelled).

44. (Cancelled)

45. (Cancelled)

46. (Cancelled)

47. (Cancelled)

48. (Cancelled)

49. (Cancelled)

**Version with Markings Illustrating Changes**

50. (Cancelled)

51. (Cancelled)

52. (Cancelled)

53. (Cancelled)

54. (Cancelled)

55. (Cancelled)

56. (Cancelled)

57. (Cancelled)

58. (Cancelled)

59. (Cancelled)

60. (Cancelled)

61. (Cancelled)

62. (Cancelled)

63. A privacy-compliant data collection and data correlation system comprising:

- a means of collecting individual-specific behavior data without knowing individual-specific demographic information pertaining to the individual about whom such data is collected;

- a means of accessing demographic data for the region in which the individual resides; and

- a means of correlating such individual-specific data with such demographic data to determine the demographic identity of each individual about whom data is collected.

64. The privacy-compliant data collection and data correlation system of Claim 63, wherein said individual-specific behavior data collection means is comprised of a set-top box.

65. (Cancelled)

66. (Cancelled)

67. (Amended) A method of reducing the effect of sampling error and sample bias on data correlations determined between a dynamic dataset and a static dataset based on assumptions about the relationships between such data, [comprising] ~~such as~~:

- creating equations to express such [data relationship] assumptions;
  - determining error functions which can assist in calculating values for each unknown variable in such equations;
  - creating a transformable matrix based on such functions;
  - inverting said matrix to apply a least-squares approach fitting method to the underlying data;
  - normalizing the results of said least-squares fit;
  - calculating Pearson-r correlations for such normalized results;
  - calculating aspect representation indices for each subset of data within said static dataset;
  - determining assumption validities for assumptions used as a basis for this process;
  - and
  - combining said correlations, said aspect representation indices, and said assumption validities to create a set of data correlations and corresponding confidence intervals.
68. The method of Claim 67 in which said dynamic dataset represents set-top box event data.
69. The method of Claim 68 in which said static dataset represents demographic information.
70. The method of Claim 69 in which the assumption used to relate said set-top box event data with said demographic information is the demographic assumption.
71. (Amended) The method of Claim [20] 67, in which said fitting procedures include applying additional assumptions to provide missing correlations values.
72. A method of increasing correlation result dataset specificity by reducing possibilities, consisting of the steps:
- calculating correlation result dataset characterization values which fall within a predetermined confidence limit using aspect representation indices, inverse demographic matrices, recombination matrices, and specification similarity matrices;
  - creating a matrix of such values for all demographic characterizations for each method used;
  - utilizing mathematical expressions of the requirement of consistency for distinct value ranges for identical characterizations in the separate matrices, reducing each range for a given characterization to the greatest possible extent within a predetermined confidence interval; thus producing one matrix with one value range for each characterization;
  - possibly transforming value ranges for all characterizations within said matrix to the same statistical confidence;

**Version with Markings Illustrating Changes**

iteratively reducing all ranges to the greatest possible extent by utilizing both mathematical expressions of the requirement of consistency among all value ranges in said matrix as well as constraints given by actual characterization population numbers; and

adjusting the statistical confidence if necessary to allow for further value range reduction past the point of useful iteration at a previous statistical confidence.

73. The method of Claim 72 in which said dataset correlations result from correlations of set-top box event data and demographic data.

74. The method of Claim 72 in which said dataset correlations result from correlations of demographic data and sales data.

75. The method of Claim 72 in which said dataset correlations result from correlations of set-top box data and sales data.

76. A method of fitting by convergence and similarity between a static dataset and a dynamic dataset, comprising the steps of:

defining subsets of each dataset;

determining correlations between such datasets;

performing a time-based analysis of group representations and additional correlations within said correlations;

assigning weights to such representations and additional correlations; and,

applying such weights and values to determine undefined correlation dataset values.

77. The method of Claim 76 in which said dynamic dataset represents set-top box data.

78. The method of Claim 77, in which said static dataset represents demographic data.

79. The method of Claim 78 in which said unidentified correlation dataset values represent non-sampled demographic specifications.

80. A method of invalidating set-top box events, comprising the steps of:

monitoring set-top box events;

storing such events in an array;

calculating trends in such events;

invalidating set-top box events which deviate in a statistically significant manner from observed set-top box event trends, or which match previously defined invalid set-top box events;

placing such invalidated set-top box events in an array; and



calculating trends in such invalidated set-top box events such that some long-term trends may be revalidated, and to identify new set-top box event categories to be ignored.